AD-A100 951 CLAREMONT MEN'S COLL CA INST OF DECISION SCIENCE F/G 12/1
MAXIMUM LIKELIHOOD ESTIMATION FOR TWO PARAMETER DECREASING FAIL--ETC(U)
JUN 81 J M MYHRE, S C SAUNDERS N00014-78-C-0213
UNCLASSIFIED 81-4 NL

1 of 1
AD A
100951

END
DATE
FILMED
7-81
DTIC

LEVEL II

12

AD A100951

INSTITUTE OF DECISION SCIENCE
FOR BUSINESS & PUBLIC POLICY

MAXIMUM LIKELIHOOD ESTIMATION FOR TWO PARAMETER

DECREASING FAILURE RATE DISTRIBUTIONS

USING CENSORED DATA

BY

Sam C. Saunders*
Washington State University
Pullman, Washington

AND

Janet M. Myhre**
Claremont Men's College
Claremont, California

Report # 81 - 4
June 1981

DTIC
ELECTE
JUN 24 1981
S D
D

DTIC FILE COPY

* Research, in part, support by Office of Naval Research Contract N00014-79-C-0755

** Research, in part, supported by Office of Naval Research Contract N000-78-C-0213

Claremont Men's College
Claremont, California

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

81 6 24 074

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 81 - 4

2. GOVT ACCESSION NO.: AD-A100951

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): Maximum likelihood estimation for two parameter decreasing failure rate distributions using censored data

5. TYPE OF REPORT & PERIOD COVERED: Technical

6. PERFORMING ORG. REPORT NUMBER: 81 - 4

7. AUTHOR(s):
Janet M./Myhre
Sam C./Saunders

8. CONTRACT OR GRANT NUMBER(s):
N00014-78-C-0213
~~N00014-79-C-0755~~

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Institute of Decision Science
Claremont Men's College
Claremont, California 91711

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Reliability

11. CONTROLLING OFFICE NAME AND ADDRESS:
Office of Naval Research
800 N. Quincy
Arlington, VA 22217

12. REPORT DATE: Jun[redacted]981

13. NUMBER OF PAGES: 41 pages

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report):

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):

N/A

18. SUPPLEMENTARY NOTES:

19. KEY WORDS (Continue on reverse side if necessary and identify by block number):

Reliability
Censored Sample
Mixed Exponential
Decreasing Failure Rate
Maximum Likelihood Estimation
Burn-in

20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

Problems of maximum likelihood estimation, for shape and scale parameters from certain decreasing hazard rate distributions which are typically mixed-exponential or work-hardened, are discussed. Sufficient conditions on the mixing distribution are given that guarantee regular behavior of the hazard rate; this ensures, even with highly censored data, that the MLE's exist whenever the sample satisfies a certain condition quite likely in samples from DHR distribution; otherwise a constant hazard rate is assumed. Some computational methods are discussed and applications made.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

410357

Maximum Likelihood Estimation for Two Parameter Decreasing Failure Rate Distribution Using Censored Data

by S. C. Saunders and J. M. Myhre

TECHNICAL SUMMARY

Let λ be the measure of the lack of resistance to sh^ck for a component, the life of which, say X_λ, will be exponential. If the variability of manufacture determines the frequency of the different λ-values, described by a r.v., $Y \sim G$, then $T = X_Y$ is the life length of a component selected at random. It will have a survival distribution

$$R(t) = E_Y P[X_Y > t|Y] = \int_0^\infty e^{-\lambda t} dG(\lambda) \quad \text{for } t > 0.$$

This is a mixed-exponential distribution with hazard $Q = -\ln R$. If G is the $\Gamma(\alpha,\beta)$ distribution we find the hazard to be $Q(t) = \alpha \ln(1 + t\beta)$, which is a Pareto type II law, known from applications in economics.

In general if $q = Q'$ is any mixed-exponential hazard rate then for $t > 0$

$$q(t) = \frac{\int_0^\infty \lambda e^{-\lambda t} dG(\lambda)}{\int_0^\infty e^{-\lambda t} dG(\lambda)},$$

and all such hazard rates are known to be decreasing.

Let $\mathfrak{Q}$ be a given class of decreasing hazard rate functions with the following properties: $q \in \mathfrak{Q}$ is twice differentiable, standardized, i.e., $q(0) = 1$, and it and the induced functions ψ and ζ, where

$$\psi(x) = xq(x) \quad \text{and} \quad \zeta(x) = 1 + xq'(x)/q(x)$$

satisfy

1^o ψ is increasing,

2^o q is log-convex,

3^o ζ has a limit at ∞ and $0 \le \zeta(\infty) \le 1$,

alternatively, sometimes the stronger condition is assumed, viz.,

$3'$ ζ is decreasing.

The unknown parameters of the life distribution under study are introduced in a manner consistent with the gamma mixed exponential; viz.,

$$R(t) = e^{-\alpha Q(t\beta)}$$

where α is the shape parameter and β the scale parameter, both positive.

Each element of this class $\mathcal{D}$ will generate a two parameter family, which is a subset of the DFR distributions.

Are there any distributions with hazard rates which satisfy these conditions? Yes, the exponential and the gamma-mixed-exponential do. Are there other DFR distributions which are not mixed-exponential which do? Yes, the failure rate of the $\Gamma(\nu,1)$ distribution does, when $0 < \nu < 1$ although it is not easy to show.

Are there any closure properties to this class? Yes, if q is a decreasing standardized hazard rate satisfying 1^o, 2^o, 3^o (or 3') then q^γ for any $0 < \gamma \leq 1$ (called an Afanas'ev generalization) does. Is $\mathcal{D}$ closed under mixtures of distributions with hazard rates in $\mathcal{D}$? Not always. If G is a distribution with support at only two points, the G-mixed exponential will, if the probability of the higher failure rate is not too large relative to the lower.

We now ask what are sufficient conditions on the mixing distribution so that assumptions 1^o, 2^o, and 3^o will be satisfied.

<u>Theorem 1</u>: If the mixing distribution G is such that the function

$$K(x,y) = G(y/x) \qquad \text{for } x,y > 0$$

is totally positive of order 2, then 1^o is satisfied, i.e., ψ_G is increasing.

We now give some conditions that a mixed hazard rate be log-convex on $(0,\infty)$.

<u>Theorem 2</u>: If for a distribution G on $(0,\infty)$ we define the convolution

$G^{(2)} = G * G$ and the related function

$$G_2(s) = \int_{u=0}^{s} \int_{t=0}^{s} (t-u)^2 \, dG(t)\, dG(u) \quad \text{for} \quad s > 0$$

then the failure rate of the G-mixed exponential distribution is log-convex if and only if for every exponential ramdom variable Z we have

$$E\, 2\Lambda_G(Z) \geq EZ\Lambda_G'(Z),$$

where

$$\Lambda_G(y) = \int_0^y s\, G_2(y-s)\, dG^{(2)}(s). \tag{1}$$

<u>Corollary 1</u>: *If the mixing distribution G is such that the induced function* Λ_G, defined in (1) is concave increasing then

$$2\Lambda_G(x) \geq x\Lambda_G'(x) \quad \text{for all } x \geq 0$$

and failure rate q_G is log-convex and 2^o is satisfied.

<u>Corollary 2</u>: *If the mixing distribution G has a density g = G' and the ratio*

$$x\, g^{(2)}(x)/g_2(x) \quad \text{for all } x > 0$$

is monotone decreasing, where g_2 and $g^{(2)}$ are the corresponding derivatives of the functions defined in theorem 2, then the failure rate q_G is log-convex.

Remark: If G is a $\Gamma(\alpha,1)$ distribution then one verifies directly that

$$\frac{x\, g^{(2)}(x)}{g_2(x)} = \frac{2\alpha + 1}{x} \qquad \text{for } x > 0$$

which decreases.

We do not consider as reasonable (for our application) mixing distributions which are of infinite order, such as $G(x) = e^{-x^{-2}}$ for $x > 0$, at the origin. We admit for consideration only distributions which are either

(i) discrete in some neighborhood of zero or

(ii) of finite order at the origin, viz., there exists $\kappa > 0$ such that $x^{-\kappa}\, G(x) \to c > 0$ as $x \to 0$.

<u>Theorem 3</u>: For any G-mixed exponential distribution the induced function ζ_G has the properties that $\zeta_G(0) = 1$, is initially decreasing, bounded above by unity, and the limit $\zeta_G(\infty)$ exists with

$$\zeta_G(\infty) = 1, \text{ if G is of type (i) and}$$

$$\zeta_G(\infty) = 0, \text{ if G is of type (ii), moreover}$$

$$\zeta_G \geq 0 \text{ if and only if } \psi_G \text{ is increasing.}$$

We now consider conditions on the mixing distribution which will insure that $1 - \zeta$ is either monotone increasing or unimodal. Such behavior can often be easily checked in specific instances, but we have a sufficient condition in

<u>Theorem 4</u>: If the associated function Λ_G, as determined from the mixing distribution G in (1), is such that the kernel

$$K(x,y) = \Lambda_G(y/x) \quad \text{for } x,y > 0$$

is totally positive of order 2, then the function $1 - \zeta_G$ induced by the failure rate of the G-mixed exponential distribution will have at most one mode in $[0,\infty]$.

It is assumed that we are given a sample vector $\underline{t} = (t_1, \ldots, t_k, \ldots, t_n)$ where $t_1, \ldots, t_k$ are ordered observations of times of failure while $t_{k+1}, \ldots, t_n$ are the ordered observed alive-times (censored lives).

The two empirical distributions of the times of failure and of the sample will be denoted by

$$F_k(y) = \frac{1}{k}\{\# \text{ of } t_i \leq y \text{ for } i = 1, \ldots, k\}$$

and similarly for F_n. We make the notational convention, to be used sub-subsequently for any function g, that its transform by an empirical distribution, is

$$\bar{g}(x) = \int_0^\infty g(xt)dF_n(t), \quad \tilde{g}(x) = \int_0^\infty g(xt)dF_k(t).$$

Thus the likelihood can be written

$$L(\alpha\ \beta|\underline{t}) = \ell n\alpha + \ell n\beta + \widetilde{\ell n q}(\beta) - \frac{\alpha n}{k}\ \bar{Q}(\beta).$$

Parenthetically, for given $\beta > 0$ the likelihood $L(\cdot|\beta, \underline{t})$ is concave on $(0,\infty)$ and the MLE of α exists uniquely and can always be obtained from the equation $L'(\alpha|\beta\cdot\underline{t}) = 0$. This tranforms all observations to an exponential with unknown failure rate α and so the MLE is given by

$$\hat{\alpha} = k/n\bar{Q}(\beta). \tag{2}$$

We examine the case for unknown β in

Theorem 5: When the shape parameter α is presumed known and a sample vector $\underline{t}$ and $q \in \mathfrak{D}$ are given, there exists an MLE of β, denoted by $\hat{\beta}$, and defined implicitly as the smallest positive root of the equation

$$\tilde{\zeta}(x) - \frac{\alpha n}{k}\ \bar{\psi}(x) = 0 \tag{3}$$

only when

$$\inf_{y>0}\ \zeta\psi^{-1}(y) < \frac{\alpha n}{k} < 1.$$

A simpler situation exists in the following case.

Theorem 6: If $q \in \mathfrak{D}$ is an Afanas'ev generalization of a G-mixed exponential, when G is of type (ii) and $0 < \gamma < 1$, with α known then for a given sample the MLE of β exists if and only if

$$1 - \gamma < \frac{\alpha n}{k} < 1.$$

Corollary 3: Under the hypothesis of theorem 6, if ζ_G is monotone then so is ζ, the corresponding function for the Afanas'ev generalization, and the MLE of β exists uniquely, i.e., there is at most only one solution to the equation (3).

We now turn to the estimation problem when both α and β are unknown.

Theorem 7: For a given sample $\underline{t}$, with $q \in \mathfrak{D}$ specified and α, β both unknown. a MLE of β, say $\hat{\beta}$, exists as the smallest positive root of the equation

$$\tilde{\zeta}(x) - \phi(x) = 0 \qquad \text{for } x > 0$$

where $\phi = \overline{\psi}/\overline{Q}$, if and only if the sample satisfies the inequality

$$2\,\overline{t}\,\tilde{t} < \overline{t^2}\ . \tag{4}$$

When $\hat{\beta}$ has been determined, the MLE of $\hat{\alpha}$, say is then given by an analogue of equation (2), namely,

$$\hat{\alpha} = k/n\ \overline{Q}(\hat{\beta})\ .$$

Our computational procedure based on the censored sample $\underline{t}$ and the assumption $q \in \mathfrak{D}$ is as follows:

Algorithm:

(a) Compute the sample moments $\tilde{t}$, $\overline{t}$, $\overline{t^2}$.

(b) If inequality (4) is not satisfied the observations are from an exponential distribution with failure rate λ and then estimate it by

$$\hat{\lambda} = \frac{k}{n\overline{t}}\ .$$

(c) If inequality (4) is satisfied then use the sample functions explicitly given by

$$\Phi(x) = x \sum_1^n t_i q(xt_i) / \sum_1^n Q(xt_i) \qquad \tilde{\zeta}(x) = 1 + \frac{x}{k} \sum_1^k [t_i q'(xt_i)/q(xt_i)].$$

We guess β_0, then iterate using the inductive step;

given β_i, compute $\Phi(\beta_i)$ and calculate β_{i+1} such that

$$\tilde{\zeta}(\beta_{i+1}) = \Phi(\beta_i).$$

We find $\hat{\beta} = \lim \beta_i$, and compute

$$\hat{\alpha} = k/\sum_1^n Q(t_i\hat{\beta}).$$

The nature of the intersection guarantees that within the region when Φ and ζ both decrease the iteration will rapidly converge, with a reasonable first guess. When the functions q and Q are simple a small programmable electronic calculator, such as the HP-67, can be used to obtain these estimates.

We now present some data sets from two different lots of flight control electronic packages. Each package has recorded, in minutes, either a failure time or an alive time, the latter is denoted by an affix +.

First Data Set

1, 8, 10, 59+, 72+, 76+, 113+, 117+, 124+, 145+, 149+, 153+, 182+, 320+.

Second Data Set

37, 53, 60+, 64+, 66+, 70+, 72+, 96+, 123+.

One checks that both data sets satisfy condition (4) so that both parameters can be estimated in a gamma-mixed-exponential model. Then using the estimation techniques derived previously we obtain:

Date Set 1	Data Set 2
$\hat{\alpha} = .0453, \hat{\beta} = 1.03$	$\hat{\alpha} = .420, \hat{\beta} = .01$

A statistical test to determine whether the data require a constant or decreasing failure rate was run on the data from sets 1 and 2. For data set 1 we reject constant failure rate (in favor of decreasing failure rate) at the .10 level. For data set 2 we cannot reject the constant failure rate assumption at that level. In this case, however, the constant failure rate estimates for reliability and the mixed exponential estimates for reliability are close.

0 Abstract

In this paper some of the problems of parametric estimation for two important types of decreasing failure rate (DFR) distributions are discussed. The first type of distribution represents the life of mixed exponential populations; the second type of distribution represents the life of mechanisms which "work harden" as they age, i.e., old components are better than new.

A representation for the second type of distribution is given so that, when satisfied, certain functional properties of the failure rate of mixed exponential distributions are guaranteed. Conditions are then imposed on the mixing distributions that will insure that the failure rate of each type of distribution satisfy these conditions.

Shape and scale parameters for any standardized DFR family of this type are to be estimated from the type of data which is available in practice, namely, severly censored samples with only a few failure observed, all of which occur early.

Conditions are obtained that maximum (rather than minimum) likelihood estimates (MLE'S) exist; the condition is phrased in terms of censored samples, namely if $t_1, \ldots, t_k$ are observed failure times, while $t_{k+1}, \ldots, t_n$ for $1 \leq k \leq n$, are censored life observations from this class of DFR distributions then the MLE's of shape and scale parameters exist if the following inequality is satisfied:

$$2 \sum_{j=1}^{k} t_j \sum_{i=1}^{n} t_i < kn \sum_{i=1}^{n} t_i^2$$

Practical methods for the computation of the MLE's are given. Actual data obtained from testing of integrated circuit electronic packages illustrate the applicability and utility of the techniques and the results described.

Key Words

Reliability
Decreasing Failure Rate
Mixed Exponential
Censored Sample
Maximum Likelihood Estimation
Burn-in

1. Introduction

Because of the demand for high reliability in electronic manufacturing, increasingly, integrated circuit modules are being used. As a consequence, modular testing now frequently encounters two conditions: the first is, the life distribution seems to exhibit a decreasing failure rate and the second is, samples are virtually always censored.

Of course these events are not unrelated, the first implies the second. The expense of testing will, when coupled with a decreasing percentage of failures over time in the surviving population, nearly always result in life length observations being censored in practice.

Several censored life data, with a paucity of failures, does not allow the usual statistical methods such as employing the Kaplan-Meier (1958) estimate of the survival function and plotting the negative of its logarithms to see if its concave or convex. In many engineering applications the monotone behavior of the hazard rate function can be deduced from the physical-chemical nature of failure. In such cases, only parameters need be estimated. There is a lack of estimation procedures which can utilize highly censored data and which avoid the potential for bias inherent in Bayesian priors.

In this paper the maximum likelihood estimates (MLE's) are obtained for both the shape and scale parameters of a class of two parameter families of decreasing failure rate (DFR) distributions and conditions are given for their existence along with practical methods of computation. These estimates are derived for censored data, which contains only a few failure observations, and *a fortiori* for complete samples.

The conditions for the existence of the MLE's apply to the case when the mean and variance of the DFR distributions do not exist. It is thought important to have estimation procedures for such distributions.

2. DECREASING FAILURE RATE MODELS

We now discuss the physical processes which determine the length of life under consideration. Firstly let us suppose the quality of construction of a component determines the level of resistance to stress which it can tolerate. Secondly, suppose the service environment provides shocks of varying magnitude to the component and failure takes place when, for the first time, the stress from an environmentally induced shock exceeds the strength of the component.

If the time between shocks exceeding any specified magnitude is exponentially distributed, with a mean depending upon that magnitude, then the life length of each component will be exponentially distributed with a failure rate which is determined by the quality of assembly. It follows that each component in service will have a constant failure rate but that the variation in manufacturing and inspection procedures will cause the population to exhibit a decreasing failure rate. Instances of such natural mixing of exponential distributions were first discussed by Proschan (1963). Subsequently other distributions with decreasing failure rates of practical interest were discussed by Cozzolino (1968). But parametric families of mized exponential and DFR distributions have received little attention compared with their IFR counterparts.

Alternatively, let us consider structures subjected to dynamic stresses of such a nature that the first stresses to which they are subjected, if not severe enough to initiate failure, only caused localized yielding and

deformation, thus effecting local stress relief and reinforcement. Such behavior increases their ability to withstand future stresses and could be thought of as "work hardening". In such cases older structures in service actually have greater resistance to fatal shocks than younger ones; i.e., each component in service has a decreasing failure rate. (An analogous behavior exists for increased strength or immunity in biological systems.) This may be thought of as "the older the better" (at least for certain periods and purposes). Usually the failure mechanism, and its interpretation in these DFR cases, is quite different from that of a mixed exponential model.

Let λ be the measure of the lack of resistance to shock for a component, the life of which, say X_λ, will be exponential. If the variability of manufacture determines the frequency of the different λ-values, which we describe by a r.v., say Λ, with distribution G, then $T=X_\Lambda$ is the life length of a component selected at random from those manufactured; it will have a survival distribution R which can be written as the conditional expectation

$$R(t) = E_\Lambda P[X_\Lambda > t|\Lambda] = \int_0^\infty e^{-\lambda t} dG(\lambda).$$

This is a mixed-exponential (survival) distribution. It is also a LaPlace transform.

One important mixed-exponential model is based on the gamma density: If for some $\alpha, \beta > 0$ we take the density

$$G'(\lambda) = \frac{\lambda^{\alpha-1} e^{-\lambda/\beta}}{\Gamma(\alpha)\ \beta^\alpha} \quad \text{for} \quad \lambda > 0 \tag{2.0}$$

then the hazard function $Q = -\ell n\ R$ is

$$Q(t) = \alpha \ln(1 + t\beta) \qquad \text{for } t > 0$$

and the density $f = Q'e^{-Q}$ is given by

$$f(t) = \frac{\alpha\beta}{(1 + t\beta)^{\alpha+1}} \quad \text{for } t > 0.$$

This two-parameter family seems to have been introduced in this country by Lomax (1954) who regarded it as a generalization of the Pareto distribution. It was called the Pareto Type II, and it has proved to be useful in business analysis.

The estimation problems for this distribution have been treated in a series of papers by Kulldorff and Vännman (1973) and Vännman (1976) but without the physical interpretation made here.

Other investigators have made use of the gamma density, as we have, to model variability among manufactured components, each one of which has a constant failure rate while in service, see the papers of Harris and Singpurwalla (1968), (1969).

In certain instances, such as in medical studies of the etiology of a disease, when the hazard rate is assumed to be either increasing or decreasing a generalization of the Pareto law, exhibiting both types of behavior depending upon its parameters, has been applied by David and Feldstein (1979). They obtained the maximum likelihood estimates implicitly for two of three parameters, neither one being a scale parameter, in the case of progressively censored data.

The question we address is how should one proceed when from physical circumstances the hazard rate is known to be mixed but by an unknown distribution?

We study the consequences of a general mixing distribution. Let q be any mixed exponential hazard rate, then for any $t > 0$

$$q(t) = \frac{\int_0^\infty \lambda e^{-\lambda t} dG(\lambda)}{\int_0^\infty e^{-\lambda t} dG(\lambda)} \tag{2.1}$$

where, without loss of generality, the scale of λ has been chosen so that q is *standardized*, i.e., $q(o) = \int_0^\infty \lambda dG(\lambda) = 1$.

The unknown parameters of the life distribution under study are introduced in a manner consistent with the gamma-mixed exponential; viz.,

$$R(t) = e^{-\alpha Q(t\beta)} \tag{2.2}$$

where α is the shape parameter and β the scale parameter, both positive.

We now introduce a model for work-hardened DFR life distributions, called an *Afanas'ev generalization*, by postulating a hazard function (for unknown $\alpha, \beta > 0$ but known γ with $0 < \gamma < 1$) of the form

$$\alpha \int_0^{t\beta} [q(x)]^\gamma dx$$

where q is defined as in (2.1) as a mixed-exponential.

It is clear that for $\gamma < 1$ we have a decreasing failure rate distribution which is not mixed exponential. In practice the value of γ is often determined by the material properties of the component in service and is not estimated by statistical techniques, see Weibull (1961).

Let us find the standardized hazard rate, say q_A, for the Afanas'ev generalization of the gamma mixture. We have

$$q_A(t) = (1 + t)^{-\gamma} \quad \text{for } t > 0 \tag{2.3}$$

from which we find the corresponding hazard function to be

$$Q_A(t) = \begin{cases} [(1 + t)^{1-\gamma} - 1]/(1 - \gamma) & \text{for } 0 < \gamma < 1 \\ \ln(1 + t) & \text{for } \gamma = 1 \end{cases} \tag{2.4}$$

For $\gamma = 1$ this is equivalent with a Weibull distribution with a location parameter of minus one and a shape parameter of less than unity. In the Soviet Union this model was introduced by Afanas'ev (1940) as a distribution for fatigue life in metals.

We now examine the distribution which results from the mixing of two exponential distributions. It is presumed that owing to occasional laxity in quality control there is a low probability $p < 1/2$ of passing a component containing a defect which can cause a high failure rate $\lambda > 1$. But there is a high probability of passing a component having the nominal (low) failure rate, which without loss of generality we take to be unity.

The reliability of this Bernoulli-mixed exponential population is

$$R(x) = pe^{-\lambda x} + (1-p)e^{-x} \text{ for } x > 0.$$

Whence, we find the standardized hazard to be, for $x > 0$

$$Q_B(x) = -\ell n p + \frac{\lambda x}{\mu} - \ell n(1 + re^{\nu x}) \tag{2.5}$$

One verifies easily that $q_B(0) = (p\lambda + q)/\mu = 1$ with

$$r = \frac{1-p}{p}, \ \nu = \frac{\lambda - 1}{\mu} > 0, \ \mu = p(\lambda + r). \tag{2.6}$$

We shall write, respectively,

$$T \sim J_A(\alpha,\beta;\gamma) \quad \text{or} \quad T \sim J_B(\alpha,\beta:p,\lambda)$$

whenever the standardized hazard Q_A, is defined in (2.4) or Q_B as defined in (2.5). Also without further mention we shall use the same subscript to denote other functions associated with these cases.

It is possible to introduce an alternative parameterization to (2.2), namely

$$R(t) = e^{-\frac{\alpha}{\beta} Q(t\beta)} \quad \text{for } t > 0; \ \alpha,\beta > 0, \tag{2.7}$$

where Q is a known cumulative hazard. The advantage of this form is that with standardized q we see from L'Hospital's rule that $\beta \to 0$ implies $R(t) \to e^{-\alpha t}$. Thus the limiting case is an exponential distribution. In either formulation one sees that β measures the rate of departure from a constant failure rate in terms of Q and hence determines the rate of decrease in the failure rate with use. The disadvantage of (2.7) is that

β is no longer a scale parameter and the likelihood equations are somewhat more complicated. However, the transformation $(\alpha,\beta) \to (\alpha/\beta,\beta)$ is a 1-1 mapping of the positive quadrant into itself so that any maximum likelihood estimates for one parameterization could be immediately transformed to the other.

3. A CLASS OF TWO PARAMETER FAMILIES OF DECREASING FAILURE RATE DISTRIBUTIONS

We shall postulate a class $\mathfrak{D}$ of concave hazard functions which has special properties encompassing both mixed exponential and "work hardened" life distributions.

Each element of this class will generate a two parameter family, which is a subset of the DFR distributions. Reliability will be of the form, where $Q \in \mathfrak{D}$ is a given hazard function,

$$R(t) = e^{-\alpha Q(t\beta)} \qquad t > 0;\ \alpha,\beta > 0.$$

Here $q = Q'$ is a decreasing hazard rate which is twice differentiable, standardized, i.e., $q(o) = 1$, and it and the induced functions ψ and ζ, where

$$\psi(x) = xq(x) \quad \text{and} \quad \zeta(x) = 1 + xq'(x)/q(x) \tag{3.1}$$

satisfy

1^o ψ is increasing

2^o q is log-convex

3^o ζ has a limit at ∞ and $0 \leq \zeta(\infty) \leq 1$,

alternatively, sometimes the stronger condition is assumed, viz.,

3' ζ is decreasing.

The question arises, "Where did such assumptions come from and what distributions, if any, satisfy them?" One sees immediately that if q decreases then q^γ decreases for any $\gamma \in (0,1)$ so that $\mathfrak{D}$ is closed under fractional powers C

of the hazard rate. The Afanas'ev generalization of the Lomax distribution has hazard rate q_A given by (2.3). One sees that q_A is decreasing and $\ln q_A$ is convex. Moreover, one checks easily that ψ_A is increasing and in this case

$$\zeta_A(t) = [1 + (1-\gamma)t]/(1+t) \tag{3.2}$$

is decreasing with $\zeta_A(\infty) = 1-\gamma$. So in this archtypical example assumptions 1^o, 2^o, and $3'$ are met.

In the Bernoulli mixture of the two exponentials we find from (2.5)

$$q_B(x) = \frac{\lambda}{\mu} - \frac{r\nu}{(e^{-\nu x}+r)}\ , \quad q_B'(x) = \frac{-\nu^2 re^{-\nu x}}{(e^{-\nu x}+r)^2} \tag{3.3}$$

and since $r > 1$,

$$q_B''(x) = \nu^3 re^{-\nu x}(r-e^{-\nu x})/(e^{-\nu x}+r)^3 > 0, \tag{3.4}$$

thus

$$\zeta_B(x) = 1 - \frac{xr\nu^2 e^{-\nu x}\mu}{(\lambda e^{-\nu x}+r)\ (e^{-\nu x}+r)}$$

so we see that $\zeta_B(\infty) = 1$ and that 3^o is satisfied.

We next show that ψ_B is an increasing function. To see this we note that $\psi_B'(x) \geq 0$ for all $x > 0$ if, after simplification,

$$\lambda e^{-\nu x} + r^2 e^{\nu x} \geq \nu^2 r\mu x - r(1 + \lambda), \text{ for all } x > 0 .$$

Let $y = e^{\nu x}$ then the inequality above becomes clearly true for $1 \leq y \leq y_o$, where $\ln y_o = (1 + \lambda)/(\lambda - 1)$. Let $y = y_o t$, for $t > 1$, then the inequality to be proved becomes

$$\frac{\lambda}{y_o t^2} + r^2 y_o \geq \nu r\mu \frac{\ell n\ t}{t}\ , \text{ for all } t > 1.$$

But the right hand side is maximized at $t = e$, so it is sufficient to have $y_o \geq \nu\mu/re$, which is implied by

$$p \leq \frac{1}{1 + (\lambda-1)e^{2/(\lambda-1)}} \simeq \frac{\lambda-1}{\lambda(\lambda + 1)} \tag{3.5}$$

(The approximation is always less than the bound.)

We now must show that ℓnq is convex. To prove this it is sufficient to show that $q_B q_B'' \geq (q_B')^2$. But substitution and simplification from (3.3) and (3.4) into the equation above shows (we spare the reader these details) that this inequality is true for all $x > 0$ iff $r \geq \sqrt{\lambda}$, if and only if

$$p \leq (1 + \sqrt{\lambda})^{-1} \tag{3.6}$$

We claim that (3.5) and (3.6) would be virtually always true in practice since if, for example, λ should be as high as 10, then the probability of passing such a bad component, with a failure rate ten times the nominal design rate, must not exceed .082. This would seem to be a reasonable assumption, at least for firms that intend to remain in business. Furthermore this demonstration shows that not all mixed exponential (and hence not all DFR) distributions satisfy our three assumptions and that there are DFR distributions, which are not mixed exponential, which do.

Let us consider the failure rate q of the gamma distribution itself, assuming a shape parameter $0 < \nu < 1$ and unit scale parameter. For $x > 0$, the failure rate satisfies the relation

$$1/q(x) = \int_0^\infty (1 + \tfrac{t}{x})^{\nu-1} e^{-t} dt = x \int_0^\infty (1 + y)^{\nu-1} e^{-y}\ dy \tag{3.7}$$

From the first equality we see $1/q$ is increasing for any $0 < \nu < 1$. From the second we see $[xq(x)]^{-1}$ is decreasing for this DFR distribution; thus 1° is satisfied. Again using the first equality in (3.7) we find that $\zeta(x)$ is non-negative, does not exceed unity and approaches one as $x \to \infty$. Hence 3° is satisfied.

To check that 2° is satisfied also is more difficult; we examine

$$\ell n q(x) = (\nu-1)\ell n\, x - x - \ell n[\int_x^{\infty} t^{\nu-1}e^{-t}dt].$$

It is thus sufficient to show that

$$(\ell n q)''(x) = \frac{1-\nu}{x^2} + q'(x) \geq 0 \quad \text{for all } x > 0. \tag{3.8}$$

From (3.7) we have, after differentiation,

$$\frac{-q'(x)}{q^2(x)} = \frac{1-\nu}{x^2} \int_0^{\infty} (1 + \frac{t}{x})^{\nu-2}\, t\, e^{-t}dt.$$

Thus to prove (3.5) it is sufficient to show that $1 - \nu \geq x^2 q'(x)$, which is the same as

$$\left[\int_0^{\infty} \left(1 + \frac{t}{x}\right)^{\nu-1} e^{-t}dt\right]^2 \geq \int_0^{\infty} \left(1 + \frac{t}{x}\right)^{\nu-2} te^{-t}dt.$$

Making the substitution $y = t/x$, then writing the squared term as the product of two integrals, one in u and the other in v, the inequality becomes

$$\int_0^{\infty} (1 + u)^{\nu-1} e^{-ux}du \int_0^{\infty} (1 + v)^{\nu-1} e^{-vx}dv \geq \int_0^{\infty} (1 + y)^{\nu-2} ye^{-yx}dy.$$

Making the change of variable $u + v = y$ on the left, the inequality becomes

$$\int_0^{\infty} \int_0^{y} (1 - u^2 + y + uy)^{\nu-1}du\, e^{-yx}dy \geq \int_0^{\infty} (1 + y)^{\nu-2} ye^{-yx}dy.$$

Thus, it is sufficient to show that the integral on the left exceeds that on the right, i.e.,

$$\int_0^y \left[1 + \frac{uy - u^2}{1 + y}\right]^{\nu-1} du \geq \frac{y}{1 + y} \quad \text{for all } 0 < u < y < \infty,$$

which is clearly true . $\square$

Of course not all DFR distributions can be obtained by an exponential mixture such as given in (2.0). For an exact description of the extreme points of the class of DFR distributions, see Langberg et al [9]. Moreover, only particular mixing distributions, thought to be of practical interest, will concern us here, along with their corresponding Afanas'ev generalizations. Thus we will examine only a subclass of the DFR distributions.

In a recent paper, McNolty, Doyle and Hansen (1980) have dealt with the mixing problem by examining some mathematical methods for inverting a general mixed-exponential reliability to obtain the mixing density. They have also discussed some of the physical interpretations of the relationship involved. We attack a related problem.

We now ask what are sufficient conditions on the mixing distribution so that assumptions 1°, 2° and 3° will be satisfied for any Afanas'ev generalizations of mixed-exponential? In the following discussion we always omit the limits of integration when they extend from 0 to ∞.

<u>Theorem 1</u>: If the mixing distribution G is such that the function

$$K(x,y) = G(y/x) \quad \text{for } x,y > 0$$

is totally positive of order 2, then 1° is satisfied, i.e., ψ_G is increasing.

Proof: Making a change of variable in the definition we see

$$\psi(t) = \int ye^{-y}dG(y/t) \Big/ \int e^{-y}dG(y/t).$$

Upon integrating numerator and denominator by parts, we obtain

$$\psi(t) + 1 = \frac{\int G(y/t)ye^{-y}dy}{\int G(y/t)e^{-y}dy} .$$

Let $t_1 > t_2$, we must show $\psi(t_1) \geq \psi(t_2)$. This is true iff

$$0 \leq \begin{vmatrix} \int G(y/t_1)ye^{-y}dy, & \int G(y/t_1)e^{-y}dy \\ \int G(y/t_2)ye^{-y}dy, & \int G(y/t_2)e^{-y}dy \end{vmatrix} .$$

By applying the basic composition formula of Karlin, see e.g., p. 100, Barlow and Proschan (1975), to the right-hand side above, it becomes equal to

$$\iint_{y_1 < y_2} \begin{vmatrix} G(y_1/t_1), & G(y_2/t_1) \\ G(y_1/t_2), & G(y_2/t_2) \end{vmatrix} \times \begin{vmatrix} y_1e^{-y_1}, & e^{-y_1} \\ y_2e^{-y_2}, & e^{-y_2} \end{vmatrix} dy_1\, dy_2.$$

Clearly the second determinant is negative and that the first is negative, by definition of TP-2, can be seen by setting $x_1 = t_2$, $x_2 = t_1$. □

The hypothesis of this theorem has a relation with Polya-Frequency functions of order 2, if it were expressed as the difference of the logarithms rather than as a ratio.

As we have previously seen for a binomial mixture of exponentials, some restrictions were necessary, namely $p \le (1+\sqrt{\lambda})^{-1}$, to insure that ℓnq was convex everywhere. We now give some conditions that a mixed failure rate be log-convex on $(0, \infty)$.

Theorem 2: If for a distribution G on $(0,\infty)$ we define the convolution $G^{(2)} = G * G$ and the related function

$$G_2(s) = \int_{u=0}^{s} \int_{t=0}^{s} (t-u)^2 \, dG(t)\, dG(u) \quad \text{for } s > 0 \qquad (3.9)$$

then the failure rate of the G-mixed exponential distribution is log-convex iff for every exponential random variable Z we have

$$E\, 2\Lambda_G(Z) \ge EZ\Lambda_G'(Z),$$

where

$$\Lambda_G(y) = \int_0^y s\, G_2(y-s)\, d\, G^{(2)}(s) \quad . \qquad (3.10)$$

Proof: From equation (2.1) we find for fixed $t > 0$ that

$$(\mu_1\mu_0)^2(\ell nq)'' = \mu_1\mu_0\,(\mu_0\mu_3-\mu_1\mu_2) - (\mu_2\mu_0-\mu_1^2)\,(\mu_2\mu_0+\mu_1^2)$$

where

$$\mu_i = \int_0^\infty \lambda^i\, e^{-\lambda t} d\, G(\lambda) \qquad \text{for } i = 0,1,\cdots .$$

Thus we may write

$$\mu_0\mu_3-\mu_1\mu_2 = \int x^3 e^{-xt} dG(x) \int e^{-yt} dG(y) - \int x^2 e^{-xt} dG(x) \int y e^{-yt} dG(y).$$

Setting, for notational simplicity,

$$d\, H(x,y) = e^{-(x+y)t}\, dG(x)dG(y) ,$$

we find it can be written

$$= \iint_{x>y} x^2(x-y)\, d\, H(x,y) + \iint_{x<y} x^2(x-y)\, d\, H(x,y).$$

and by making an interchange of variables in the second integral, obtain

$$\mu_0\mu_3 - \mu_1\mu_2 = \iint_{x>y} (x+y)(x-y)^2 dH(x,y).$$

By a similar argument we obtain

$$\mu_0\mu_2 - \mu_1^2 = \iint_{x>y} (x-y)^2 \, dH(x,y).$$

From arguments of symmetry we see that

$$\mu_1\mu_0 = \iint x \, dH(x,y) = \iint y \, dH(x,y)$$

and so

$$\mu_0\mu_1+\mu_1^2 = \iint (x^2+xy) \, dH(x,y) = \iint (y^2+xy) \, dH(x,y).$$

Hence we see that for $c = 2(\mu_1\mu_0)^2$, with the obvious change of variable

$$c(\ell nq)'' = \iint (u+v) \, dH(u,v) \iint_{x>y} (x+y)(x-y)^2 \, dH(x,y)$$
$$- \iint (u+v)^2 \, dH(u,v) \iint (x-y)^2 \, H(x,y).$$

Now let $w = u+v$, and simplify to find that

$$2c(\ell nq) = \int_0^\infty [\iint (x-y)^2 (x+y-w) \, dH(x,y)] \, w \, e^{-tw} dG^{(2)}(w).$$

The quantity in square brackets above can be written

$$[. \; . \; .] = \int_{x=0}^\infty \int_{y=0}^\infty (x-y)^2(x+y-w)e^{-t(x+y)} dG(y)dG(x).$$

Letting $x + y = s$ we obtain, in the case the density $G'=g$ exists,

$$[. \; . \; .] = \int_0^\infty \int_{y=0}^\infty (s-2y)^2(s-w)e^{-ts} g(s-y) \, g(y)dyds.$$

For this case we set

$$g_2(s) = \int_0^0 (s-2y)^2 g(s-y) \, g(y) \, dy \qquad \text{for } s>0,$$

and verify that $G_2' = g_2$. In the general case the quantity becomes

$$[. \; . \; .] = \int_0^\infty (s-w)e^{-ts} dG_2(s)$$

Hence for some $c_1 > 0$ we have

$$c_1(\ell nq)'' = \int_0^\infty \int_0^\infty w(s-w)\, e^{-t(s+w)}\, dG_2(s)\, dG^{(2)}(w). \qquad (3.11)$$

Assuming, only for notational convenience, that both g_2 and $g^{(2)}$ exist we have

$$c_1(\ell nq)'' = \int_0^\infty e^{-tx} \left\{ \int_0^x (x-s)(2s-x)\, g_2(s)\, g^{(2)}(x-s)\, ds \right\} dx.$$

Let us set

$$f(s,x) = s\, g_2(x-s)\, g^{(2)}(s) \qquad \text{for } 0 \le s \le x, \qquad (3.12)$$

then we can rewrite the quantity in braces above, after breaking the integral into two parts and changing variables in the second, as

$$\{.\ .\ .\} = \int_0^{x/2} (x-2s)[f(s,x)-f(x-s,x)]ds = 2\Lambda_G(x) - x\Lambda_G'(x).$$

The second equality is obtained by integrating each term of the difference by parts and simplifying the resulting expressions where

$$\Lambda_G(y) = \int_0^y \int_0^x s g_2(x-s) g^{(2)}(s)\, ds dx.$$

By utilizing the Liebniz rule for change of the order of integration this can be seen to be equal to the expression for Λ_G given in the hypothesis: The proof of the theorem is completed upon noting that

$$c_1(\ell nq)'' = \int_0^\infty e^{-tx}[2\Lambda_G(x)-x\Lambda_G'(x)]\, dx\ .\ \Box \qquad (3.13)$$

We now give some sufficient conditions that ℓnq is convex.

Corollary 1: If the mixing distribution G is such that the induced function Λ_G, defined in (3.10) is concave increasing then

$$2\Lambda_G(x) \geq x\Lambda_G'(x) \quad \text{for all } x \geq 0$$

and the failure rate q_G is *log-convex*.

The proof is obvious since Λ_G' is positive, decreasing.

Corollary 2: If the mixing distribution G has a density $g = G'$ and the ratio

$$x\, g^{(2)}(x)/g_2(x) \qquad \text{for } x > 0$$

is monotone decreasing, where g_2 and $g^{(2)}$ are the corresponding derivatives of the functions defined in theorem 2, then the failure rate q_G is log-convex.

Proof: It is sufficient to note that

$$f(s,x) \geq f(x-s,x) \qquad 0 < s < x/2,$$

where f was defined in (3.12). Thus we must show

$$s g^{(2)}(s)\, g_2(x-s) \geq (x-s) g_2(s)\, g^{(2)}(x-s) \qquad 0 < s < x/2,$$

which upon division is seen to be guaranteed by the monotone behavior of the ratio assumed in the hypothesis . □

Remark: If G is a $\Gamma(\alpha, 1)$ distribution then one verifies directly that

$$\frac{x\, g^{(2)}(x)}{g_2(x)} = \frac{2\alpha + 1}{x} \qquad \text{for } x > 0$$

which decreases.

By Corollary 2 we are assured that the gamma mixture of exponentials will have a failure rate which is log-convex: A fact which can be verified directly from (1.4).

We now examine the behavior of the induced function ζ where $Q \in \mathcal{Q}$. We note that always, since $q' \leq 0$ and $\psi' \geq 0$, that

$$1 \geq 1 + \frac{xq'}{q} = \zeta = \frac{\psi'}{q} \geq 0$$

If ζ is a given function bounded between zero and one, differentiable at zero with $\zeta(0) = 1$ and $\zeta'(0) \leq 0$, having a limit at ∞ and for which the ratio $\frac{\zeta-1}{t}$ is increasing for $t > 0$ then by regarding the expression (3.1.B) as a differential equation in the unknown function q the solution

$$q(t) = \exp\{-\int_0^\tau \frac{1-\zeta(x)}{x}\,dx\} \qquad \text{for } t > 0$$

defines a decreasing failure rate, which is log-convex, standardized and for which $t\,q(t)$ is increasing, i.e., it satisfies the assumptions 1^o, 2^o and 3^o. We note that $\zeta(t) = 1-\gamma + \gamma(1+t)^{-1}$ generates the Afanas'ev generalization of the gamma-mixed exponential model. We now study the behavior of the induced ζ in the case of a general mixed exponential, in particular its asymptotic behavior as determined by the behavior of the appropriate mixing distribution G near 0.

We do not consider as reasonable (for our application) mixing distributions which are of infinite order at the origin, such as $G(x) = e^{-x^{-2}}$ for $x > 0$. We admit for consideration only distributions which are either

(i) discrete in some neighborhood of zero or

(ii) of finite order at the origin, viz., there exists $\kappa > 0$ such that $x^{-\kappa} G(x) \to c > 0$ as $x \to 0$

Theorem 3: For any G-mixed exponential distribution the induced function ζ_G has the properties that $\zeta_G(0) = 1$, is initially decreasing, bounded above by unity, and the limit $\zeta_G(\infty)$ exists with

$$\zeta_G(\infty) = 1, \text{ if } G \text{ is of type (i) and}$$
$$\zeta_G(\infty) = 0, \text{ if } G \text{ is of type (ii), moreover}$$
$$\zeta_G \geq 0 \text{ iff } \psi_G \text{ is increasing.}$$

Proof: Considering equation (2.1), we can write for any $t > 0$,

$$1 - \zeta(t) = \frac{-tq(t)}{q(t)} = \frac{t\int e^{-xt} d\, G_2(x)}{\int y\, e^{-yt} d\, G^{(2)}(y)} . \qquad (3.14)$$

using the methods and notation of theorem 3 .

Thus we see from (3.14) that $\zeta(o) = 1$, and since the right hand side is positive, that $1 \geq \zeta(t)$. We see ζ initially decreases linearly with slope $q'(o)$ since

$$\zeta'(o) = \lim_{x \downarrow o} \frac{\zeta(x)-1}{x} = q'(o).$$

Making use of classical Tauberian theorems on LaPlace transforms, e.g., Widder (1946) p. 181, we obtain the limiting behavior of ζ_G at ∞.

The last claim follows from the identity $\zeta_G = \psi_G'/q_G$. □

We now consider conditions on the mixing distribution which will insure that $1 - \zeta$ is either monotone increasing or unimodal. Such behavior can often be easily checked in each specific instance but we have a sufficient condition in

Theorem 4: If the associated function Λ_G, as determined from the mixing distribution G in (3.10), is such that the kernel

$$K(x,y) = \Lambda_G(y/x) \quad \text{for } x,y > 0$$

is totally positive of order 2, then the function $1 - \zeta_G$ induced by the failure rate of the G-mixed exponential distribution will have at most one mode in $[0, \infty]$

Proof: By taking the derivative of equation (3.14) and considering only numerator we find, utilizing notation from the hypothesis of theorem 2, that

$$\text{num}[-\zeta_G'(t)] = \iint [-ty(x-y) + y]\, e^{-(x+y)t} \, dG^{(2)}(y)\, dG_2(x).$$

By comparing the first term above with equation (3.11) and using the representation in (3.13), as well as making a change of variable in the second term to obtain a convolution, we find after simplification

$$= \int (y-2)e^{-y} \quad \Lambda_G(y/t)\, dy.$$

We note the function $(y-2)e^{-y}$ for y 0 changes sign exactly once, and so by applying the variation diminishing properties of Polya frequency functions, see e.g., p. 93, Barlow and Proschan, we conclude that $-\zeta_G'(t)$ changes sign at most once. Therefore, $1 - \zeta_G$ possesses at most one mode. □

In any life length model one is interested in the distribution resulting when independent components, from the same family, are connected in a series system. From the concavity of the hazard $Q \in \mathfrak{Q}$ follows the

Remark: If components with independent life lengths $T_i \sim J_Q(\alpha_i, \beta_i)$ for $i = 1,\ldots,n$, for some $Q \in \mathfrak{Q}$, are in series then the life of the system T will satisfy the stochastic inequality

$$T = \min T_i \gtrsim J_Q\,(\Sigma\alpha_i,\ \Sigma\alpha_i\beta_i/\Sigma\alpha_i)$$

with equality when $\beta_i \equiv \beta$.

Another property one would ask of any model in which the failure rate is initially decreasing is what improvement can be made by "burning-in" a component having such a life? In practice it is often assumed that as a result of a burn-in period, surviving components are exponentially lived. In fact, burn-in tests are often required in electronic component procurement with a statement of the ultimate failure rate so obtained.

Of course, not all decreasing failure rate distributions do become constant after some finite initial period, but that is an assumption which is often thought to be appropriate. This indicates the importance of the second model introduced and of its utility in a determination of the economic value of the stochastically extended life.

The residual life T_τ of a component with new life length T and a burn-in of duration $\tau > 0$ is the conditional life remaining after time τ given that it is alive then; i.e., $T_\tau = [T-\tau | T > \tau]$.

The residual life of any G-mixed exponential is again a mixed exponential but with a different mixing distribution. The residual life will have density $f(y) = ce^{-\tau y} dG(Y)$, where c is the normalizing constant.

But is the class $\mathfrak{Q}$ closed under burn-in? A burn-in of duration τ for a component with hazard $Q \in \mathfrak{Q}$ will yield a residual life T_τ with hazard rate

$$q_\tau(t) = q\,(t + \tau) \quad \text{for } t > 0.$$

Clearly q_τ is not standardized but one sees, after brief reflection, that ψ_τ is increasing, q_τ is log-convex, ζ_τ approaches a limit in [0,1], if ζ does, and lastly, ζ_τ is decreasing if ζ is. Thus the answer is affirmative, except for standardization.

It is easily seen that T_τ is stochastically larger than T for all $\tau > 0$ but usually the burn-in time τ is increased until any incremental decrease in the

residual failure rate is not worth the incremental cost of testing. This point will necessarily be somewhat different for each particular Q.

Remark: A burn-in of τ units of time on a component with a new life $T \sim J_A(\alpha,\beta)$ will yield a residual life

$$T_\tau \sim J_A[\alpha(1+\tau\beta)^{1-\gamma},\ \beta/(1+\tau\beta)].$$

If $T \sim J_B(\alpha,\beta:p,\lambda)$ then a burn-in length τ will only alter the proportion of high failure rates, viz.,

$$T_\tau \sim J_B(\alpha,\beta:p',\lambda)$$

where the altered proportion is given

$$p' = pe^{-\lambda\tau\beta}/(pe^{-\lambda\tau\beta}+(1-p)e^{-\tau\beta}).$$

4. ESTIMATION OF PARAMETERS USING INCOMPLETE SAMPLES

The samples, that are obtained when components having a DFR life distribution are tested, are virtually always incomplete in the sense that testing is stopped before all components have failed. A datum on a component that "failure has not yet occurred after a specified life" is called in practice an alive time or a run-out. Samples containing such observations are said to be *censored*. Samples in which life tests are truncated at some preassigned ordered observation occur infrequently, in our experience, when electronic components are tested.

It is assumed throughout this section that we are given a sample vector $\underline{t} = (t_1, \ldots, t_k, \ldots, t_n)$ where $t_1, \ldots, t_k$ are ordered observations of times of failure while $t_{k+1}, \ldots, t_n$ are the ordered observed alive-times, with

$1 \le k \le n$. All observations are presumed to have been obtained by testing components having a $J_Q(\alpha,\beta)$ distribution with unknown parameters α and β, but with $Q \in \mathfrak{Q}$ and γ specified.

We now introduce notation for the two empirical distributions (call them F_k and F_n) of the times of failure and of the sample, respectively. We set

$$F_k(y) = \frac{1}{k}\{\# \text{ of } t_i \le y \text{ for } i = 1, \ldots, k\}$$

and similarly for F_n, and we make the notational convention, to be used subsequently for any function g, that its transform, according to the empirical distribution, is denoted by the proper affix,

$$\bar{g}(x) = \int_0^\infty g(xt)dF_n(t), \quad \tilde{g}(x) = \int_0^\infty g(xt)dF_k(t).$$

Some results will now be given on maximum likelihood estimation of the unknown shape and scale parameters in the case of censored samples with $Q \in \mathfrak{Q}$ and $\gamma \in (0,1]$ specified.

The sample vector

$$\underline{t} = (t_1, \ldots, t_k, \ldots t_n) \text{ for } 1 \le k \le n$$

corresponds to the observed events

$$[T_i = t_i] \text{ for } i = 1, \ldots, k \text{ and } [T_i > t_i] \text{ for } i = k+1, \ldots, n.$$

By definition the log-likelihood, after substituting from (2.1), is given by

$$k \,\ell n(\alpha\beta) + \sum_{i=1}^{k} \ell n\, q(t_i\beta) - \alpha \sum_{i=1}^{n} Q(t_i\beta).$$

Dividing by the constant k, we write the resulting function of the parameters α,β, given the vector $\underline{t}$, as

$$L(\alpha, \beta|\underline{t}) = \ell n\, \alpha + \ell n\, \beta + \widetilde{\ell n q}(\beta) - \frac{\alpha n}{k} \bar{Q}(\beta), \tag{4.1}$$

where we have made use of the notational convention introduced earlier.

Parenthetically, for given $\beta > 0$ the likelihood $L(\cdot|\beta, \underline{t})$ is concave on $(0, \infty)$ and the MLE of α exists uniquely and can always be obtained from the equation $L'(\alpha|\beta.\underline{t}) = 0$.

We thus have

Remark: When the scale parameter $\beta > 0$ is given, there exists a unique MLE of α, say $\hat{\alpha}$, given explicitly by

$$\hat{\alpha} = k/n\, \bar{Q}\, (\beta). \tag{4.2}$$

This result is well known. If β is known and Q given, concave or not, then the values $y_i = Q(t_i\beta)$ for $i=1, \ldots, n$ can be calculated. They are the alive and dead times from an exponential distribution with unknown failure rate α. The total life statistic divided by the number of failure yields the usual maximum likelihood estimate of the mean life.

It is also true that not any set of n positive numbers $(t_1, \ldots, t_n)$ with $1 \leq k \leq n$ designated as failure times and the remainder as alive times can be used to estimate uniquely both the unknown parameters for any $Q \varepsilon \mathfrak{D}$. In some sense the sample must be close to what would be likely from such a hazard function.

We examine the case for unknown β in

Theorem 5: When the shape parameter α is presumed known and a sample vector $\underline{t} = (t_1, \ldots, t_k, t_{k+1}, \ldots, t_n)$ and $Q \in \mathfrak{D}$ are given, there exists an MLE of β, denoted by $\hat{\beta}$, and defined implicitly as the smallest positive root of the equation

$$\tilde{\zeta}\,(x) - \frac{\alpha n}{k}\, \bar{\Psi}\,(x) = 0 \tag{4.3}$$

only when

$$\inf_{y > 0} \zeta\psi^{-1}(y) < \frac{\alpha n}{k} < 1 \tag{4.4}$$

Remark: The condition (4.4) is determined principally by the behavior of ζ, since ψ is always increasing by assumption 2^o, mapping $(0, \infty)$ into $(0, \infty)$. For example, ζ may be monotone decreasing, as is ζ_A with $\zeta_A(\infty) = 1-\gamma$, while ζ_B only decreases initially but is concave-increasing, ultimately, with $\zeta_B(\infty) = 1$.

Proof of theorem 6: From (4.1) we see the likelihood $L(\beta|\alpha, \underline{t})$ can be written, neglecting constants not depending upon β, as

$$L(\beta|\alpha, \underline{t}) = \widetilde{\ln\psi}(\beta) - \frac{\alpha n}{k} \bar{Q}(\beta)$$

If we define $A(\beta) = L'(\beta|\alpha, \underline{t})$, we see we must determine the roots of

$$A(x) = \tilde{\zeta}(x) - \frac{\alpha n}{k} \bar{\psi}(x)$$

which confirms equation (4.3) as the appropriate one. To check the necessary condition we see the likelihood increases initially as x increases since $A(0) = \tilde{\zeta}(0) = 1$. But $\bar{\psi}$ is always increasing, while $\tilde{\zeta}$ is only initially decreasing, since

$$\lim_{x \downarrow 0} \tilde{\zeta}'(0) = \zeta'(0)\,\tilde{t} < 0.$$

But, by assumption 3^o, the limit $\zeta(\infty)$ is in the unit interval. Thus for a solution to the equation

$$\tilde{\zeta}\,[(\bar{\psi})^{-1}(y)] = \frac{\alpha n}{k} \qquad \text{for } y > 0$$

to exist, the range of the composed function must contain the value $\alpha n/k$, as stated in the hypothesis. □

We see that the second smallest positive root of equation (4.2), if it exists, will be a minimum likelihood estimate of β. This situation can occur frequently. If there are either more than one local maxima to $L(\beta|\alpha, \underline{t})$, i.e., three or more roots to (4.2), or there is no maximum in $(0,\infty)$, then this would indicate that the presupposed choice of α and/or of $Q \in \mathfrak{D}$ should be reexamined. That is to say, either the initial failure rate α is not of the right magnitude to reflect the number of first failures observed or the induced function ζ does not decrease monotonely over a sufficiently long interval.

A simpler situation exists in the following case.

Theorem 6: If $Q \in \mathfrak{D}$ is an Afanas'ev generaliztion of a G-mixed exponential, when G is of type (ii) and $0 < \nu < 1$, with α known then for a given sample the MLE of β exists iff

$$1 - \gamma < \frac{\alpha n}{k} < 1 .$$

Proof: We note by theorem 3, since G is of type (ii), that the range of ζ_G, the induced function for the G-mixed exponential, is [0,1]. Letting $q = (q_G)^\nu$ we find $\zeta = 1-\gamma + \gamma\zeta_G$ so that the range of ζ is $[1-\nu, 1]$. We find from equation (2.1) that

$$\psi(x) = x^{1-\nu} \left[\frac{\int G(y/x)d(ye^{-y})}{\int G(y/x)d(e^{-y})} \right]^\gamma$$

so that $\lim_{x\to\infty} \psi(x) = \infty$. thus by equation (4.4) the result follows. $\square$

Corollary 3: Under the hypothesis of theorem 7, if ζ_G is monotone then so is ζ and the MLE of β exists uniquely, i.e., there is at most only one solution to the equation (4.3).

We now turn to the estimation problem when both α and β are unknown.

Theorem 7: For a given sample $\underline{t}$, with $Q \in \mathfrak{D}$ specified and α,β both unknown, a MLE of β, say $\hat{\beta}$, exists as the smallest positive root of the equation

$$\tilde{\zeta}(x) - \phi(x) = 0 \qquad \text{for } x > 0$$

where $\phi = \bar{\Psi}/\bar{Q}$, iff the sample satisfies the inequality

$$2\,\bar{t}\,\tilde{t} < \overline{t^2} \tag{4.5}$$

When $\hat{\beta}$ has been determined, the MLE of $\hat{\alpha}$, say is then given by an analogue of equation (4.2), namely,

$$\hat{\alpha} = k/n\,\bar{Q}(\hat{\beta}). \tag{4.6}$$

Proof of theorem 7: Consider the likelihood function $L(\alpha,\beta|\underline{t})$, as given in (4.1), defined over the positive quadrant. All the stationary points, which are determined by $\underline{t}$, can be found from the simultaneous solution of $\frac{\partial L}{\partial \alpha} = 0$, $\frac{\partial L}{\partial \beta} = 0$. This yields the two equations in α and β ;

$$\tilde{\zeta}(\beta) = \frac{\alpha n}{k}\,\bar{\Psi}(\beta); \quad \frac{1}{\alpha} = \frac{n}{k}\,\bar{Q}(\beta).$$

Combining these into a single equation in the unknown β, we are led to seek $\hat{\beta}$ as a zero of the difference $\tilde{\zeta}(x) - \Phi(x)$ for $x > 0$ where $\Phi = \bar{\Psi}/\bar{Q}$. If we maximize the likelihood with respect to α for any value of β by substituting $\alpha = k/n\bar{Q}(\beta)$ into (4.1) and obtain $L(k/nQ(\beta),\beta|\ \underline{t})$, then neglect constants independent of β, we obtain a function, call it B, which can be written, now using argument x, instead of β, as

$$B(x) = \widetilde{\ell n q}(x) - \ell n[\overline{Q}(x)/x].$$

Interchanging the order of integration in $\overline{Q}(x)/x$ we obtain

$$\overline{Q}(x)/x = \overline{t}\int_0^\infty q(tx)\, dW_n(t) \equiv \overline{t}\cdot q^*(x)$$

where W_n is the distribution, with density given by

$$W_n'(t) = [1 - F_n(t)]/\overline{t} \qquad \text{for } t > 0.$$

Again neglecting constants, we have

$$B(x) = \int_0^\infty \ell n q(tx)\, dF_k(t) - \ell n\left[\int_0^\infty q(tx)\, dW_n(t)\right]$$

$$= \widetilde{\ell n q}(x) - \ell n q^*(x) \; . \tag{4.7}$$

Using a Maclaurin expansion of q we see

$$B(x) = x\, q'(0)\, [\tilde{t} - t^*] + O(x^2) \qquad \text{as } x \to 0,$$

where

$$\tilde{t} = \int t\, dF_k(t) \quad \text{and} \quad t^* = \int t\, dW_n(t) = \overline{t^2}/2\overline{t} \; .$$

Because $q'(0) < 0$, it follows that B is positive in a neighborhood of zero iff $\tilde{t} < t^*$. Since $B(0) = B(\infty) = 0$, we must have $B'(x) = 0$ for some $x > 0$. Moreover, one verifies that $xB' = \tilde{\zeta} - \Phi$ and so the MLE can be found as the smallest positive root of this difference. (The second smallest root, if it exists, will be a minimum likelihood estimate.) Having determined $\hat{\beta}$ one uses the partial derivative equated to zero to obtain $\hat{\alpha}$. □

We now discuss the situation when the sample fails to satisfy the condition $2\overline{t}\,\tilde{t} < \overline{t^2}$, and MLE's do not exist uniquely. This means that the model, i.e., the choice of Q, may not be appropriate and either a constant failure rate model or a convex failure rate model (one that is initially decreasing and ultimately increasing) may be indicated rather than a DFR model with $Q \in \mathfrak{D}$.

5. THE COMPUTATION OF $\hat{\beta}$ FOR CENSORED SAMPLES

As a matter of practical calculation we are concerned with the smallest root of the difference $\tilde{\zeta} - \Phi$, in the case when ζ is monotone decreasing. Equivalently, let us consider the composite function $f(x) = \tilde{\zeta}^{-1}[\Phi(x)]$ in a neighborhood of zero with the location of the smallest crossing, if there is more than one, of the 45° line.

An alternative expression for Φ is $\Phi = x\bar{Q}'/\bar{Q}$, with $\bar{Q}'$ a convex function decreasing between $\bar{t}$ and $\bar{t} \cdot q(\infty)$, while the smoothed $\bar{Q}'$, viz., $\bar{Q}/x$, decreases between the same limits at a slower rate. It follows that $\bar{Q}/x \geq \bar{Q}' \geq 0$ so that $0 \leq \Phi \leq 1$. Then Φ begins at unity, initially decreases at a decelerating rate and tends ultimately to $\zeta(\infty)$. To see this note

$$\Phi(\infty) = \lim_{\infty} \bar{\psi}' / \bar{Q}' = \lim_{x\to\infty} \frac{\int t^{1-\nu}(xt)^{\nu} q(tx)\zeta(tx)\,dF_n(t)}{\int t^{1-\nu}(xt)^{\nu} q(tx)\,dF_n(t)} = \zeta(\infty)$$

since by Widder loc. cit. there exists a $\upsilon \in [0,1]$ such that $x^{\upsilon}q(x) \to a \neq 0$, as $x \to \infty$. The composite function $f = \tilde{\zeta}^{-1}\Phi$ behaves in a neighborhood of zero as a contractive map, being initially greater than x, then crossing at $\hat{\beta}$ and then being below x for a range, perhaps, thereafter.

Thus we know that successive iterates

$$\beta_{i+1} = f(\beta_i) \qquad \text{for } i = 0,1,2 \ldots$$

will converge to $\beta > 0$ as long as $\beta_0 < \beta^{0}$, the next larger zero of $\tilde{\zeta} - \Phi$, if one exists. Otherwise the iteration will converge to zero.

Moreover for the special cases in (3.2) the inverse $\tilde{\zeta}^{-1}$ can be easily found.

Our computational procedure based on the sample is as follows:

Algorithm: Given $t_1, \ldots, t_k$ as failure times and $t_{k+1}, \ldots, t_n$ as censored live times from a DFR distribution, proceed as follows:

(i) Compute the sample moments $\tilde{t}, \bar{t}, \overline{t^2}$.

(ii) If $\overline{t^2} < 2\bar{t} \cdot \hat{t}$, assume the observations are from an exponential distribution with failure rate λ and estimate it by

$$\hat{\lambda} = \frac{k}{n\bar{t}} .$$

(iii) If $\overline{t^2} > 2\bar{t}\,\tilde{t}$, assume the observations are from a DFR distribution, with prescribed $Q \in \mathfrak{Q}$.

Using the sample functions explicitly given by

$$\phi(x) = x \sum_1^n t_i q(xt_i) / \sum_1^n Q(xt_i) \qquad \tilde{\zeta}(x) = 1 + \frac{x}{k} \sum_1^k [t_i q'(xt_i)/q(xt_i)],$$

we guess β_0, then iterate using the inductive step;

given β_i, compute $\Phi(\beta_i)$ and calculate β_{i+1} such that

$$\tilde{\zeta}(\beta_{i+1}) = \Phi(\beta_i).$$

we find $\hat{\beta} = \lim \beta_i$, and compute

$$\hat{\alpha} = k / \sum_1^n Q(t_i \hat{\beta}).$$

The nature of the intersection guarantees that within the region when Φ and $\tilde{\zeta}$ both decrease the iteration will rapidly converge, with a reasonable first guess. When the functions q and Q are simple a small programmable electronic calculator, such as the HP-67, can be used to obtain these estimates.

We now present some data sets from two different lots of flight control electronic packages. Each package has recorded, in minutes, either a failure time or an alive time. An alive time is the time the life test was terminated with the package still functioning and is denoted by an affix +.

First Data Set

1, 8, 10, 59+, 72+, 76+, 113+, 117+, 124+, 145+, 149+, 153+, 182+, 320+.

Second Data Set

37, 53, 60+, 64+, 66+, 70+, 72+, 96+, 123+.

One checks that both data sets satisfy condition (4.5) so that both parameters can be estimated in a gamma mixed-exponential model. Then using the estimation techniques derived previously in this paper we have the following estimates:

Data Set 1	Data Set 2
$\hat{\alpha} = .0453, \hat{\beta} = 1.03$	$\hat{\alpha} = .420, \hat{\beta} = .01$

A statistical test to determine whether the data require a constant or decreasing failure rate was run on the data from sets 1 and 2. For data set 1 we reject constant failure rate (in favor of decreasing failure rate) at the .10 level. For data set we cannot reject the constant failure rate assumption. In this case, however, the constant failure rate estimates for reliability and the mixed exponential estimates for reliability are close.

7. Conclusion

We are primarily concerned with the problem of estimating the hazard rate of a component, such as an integrated circuit, assuming that it has either a constant hazard rate or a decreasing hazard rate of specified functional form with shape and scale parameters unknown. The influence of the data is different in this case than in the more usual case of hazard rate known to be increasing.

If a component has a life distribution with an increasing failure rate, the information necessary to estimate its parameters must contain failure times. In practice this means that if there are few observed failures, within a fleet of orperational components, there is little information with which to assess their reliability. If a component has a constant failure rate then both failure times and alive times contribute equally to the estimation of its reliability. This study suggests that if a component has a life distribution with decreasing failure rate it is the alive times within the data which contribute principally to the estimation of the parameters (and thereby to the determination of the reliability) since only one failure observation is required even to estimate two parameters, presuming the data is ample.

Note that for a sample of size two, both of which are failure observations, the inequality (4.5) cannot be satisfied since $t_1 t_2 > 0$ implies that $2(\bar{t})^2 > (t_1^2 + t_2^2)/2$; but if t_1 is a failure while t_2 is an alive time, for which $(1 + \sqrt{2})t_1 < t_2$, then the inequality is true.

To illustrate the typical behavior of the likelihood B, as given in equation (4.7), let us assume that the sample distributions F_k and W_n as defined previously satisfy $F_k \geq W_n$; this implies $\tilde{t} \geq t^*$. (Strictly speaking this condition cannot ever be met since even if the failure times are stochastically smaller than the sample containing failure and alive times

together, i.e. $F_k \geq F_n$, still we would have $0 = F_k < W_n$ on $(0,t_1)$. But since $\overline{F}_k \leq \overline{W}_n$ on (t_1,∞) we would have $\tilde{t} \leq t^*$ if and only if

$$\int_{t_1}^{\infty} [\overline{W}_n(t) - \overline{F}_k(t)] \; dt \geq t_1^2 \; [1 - 1/2n \; \overline{t}] \; .$$

Because t_1 is very small in practice this is virtually always true.) From this assumption would follow $\tilde{q} \geq q^*$ and since ℓn is concave, that $\ell n \; q^* \geq (\ell n \; q)^*$. Hence from (4.7) one sees that

$$\widetilde{\ell n \; q} - \ell n \; \tilde{q} \leq B \leq \widetilde{\ell n \; q} - (\ell n \; q)^*.$$

Now one notices that the lower bound is alwasy negative while the upper bound is always positive. But near the origin B is approximately equal to the upper bound while for large values of its argument B is approximately equal to the lower bound and for intermediate values B makes a transition between.

Thus for any sample in which the failure times are stochastically smaller than the combined sample times and for which $2\overline{t} \; \tilde{t} < \overline{t^2}$ we find a maximum likelihood estimate of β as the smallest positive root and a minimum likelihood estimate of β as the second smallest root. In the case $2\overline{t} \; \tilde{t} > \overline{t^2}$, which could occur with complete failure data, we would have a minimum likelihood estimate as the smallest positive root and a maximum likelihood estimate at $\hat{\beta} = 0$. The frequency with which this occurred under various mixtures was studied by Sunjata (1974). Maxima occurring at such boundaries were also observed by Davis and Feldstein (1979) in their study. Of course by arbitrarily grouping failures several local extrema of B can be constructed. This is regarded as being of little practical significance.

The usual justification for using maximum likelihood estimates is due to their asymptotically optimal properties, and to their asymptotic normality. The problem of obtaining the usual sampling distributions of the MLE's of the parameters obtained for these DFR models seems to be difficult, not only

because the estimates are only implicitly defined, but also because samples are virtually always censored. Furthermore, the usual proofs for the asymptotic optimality of the MLE's may not apply when censoring is of a general type and when only sparse failure data are available. A useful asymptotic theory must be developed for censored data sets of which the life of electronic pacages of integrated circuits are an illustration.

REFERENCES

Afanas'ev, N. N. (1940). Statistical theory of the fatique strength of metals. *Zhurnal Tekhnicheska Fiziki, 10,* 1553-1568.

Barlow, R. E., and Proschan, Frank (1975). *Statistical Theory of Reliability and Life Testing.* Holt, Rinehart and Winston, Inc., New York.

Cozzolino, John M. (1968). Probabilistic models of decreasing failure rate processes. *Naval Research Logistic Quarterly, 15,* 361-374.

Davis, H. T., and Feldstein, M. L. (1979). The generalized Pareto law as a model for progressively censored survival data. *Biometrika, 66,* 299-306.

Harris, C. M., and Singpurwalla, N. D. (1968). Life distributions derived from stochastic hazard functions. *IEEE Trans. in Reliability, R-17,* 70-79.

Harris, C. M., and Singpurwalla, N. D. (1969). On estimation in Weibull distributions with random scale parameters. *Naval Research Logistic Quarterly, 16,* 405-410.

Kaplan, E. L., and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Stat. Assoc., 53,* 457-481.

Kulldorff, G., and Vännman, K. (1973). Estimaton of the location and scale parameters of a Pareto distribution. *J. Amer. Statist. Assoc., 68,* 218-227.

Langberg, Naftali A., Lion Ramon V., Proschan, Frank, and Lynch, James (November 1978). The extreme points of the set of decreasing failure rate distribution. FSU Technical Report.

Lomax, K. S. (1954). Business failures: Another example of the analysis of failure data. *J. Amer. Statist. Assoc., 49,* 847-852.

McNolty, F., Doyle, J., and Hansen, E. (1980). Properties of the mixed exponential failure process. *Technometrics, 22,* 555-566.

Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics, 5,* 375-383.

Sunjata, M. H. (1974). *Sensitivity Analysis of a Reliability Estimation Procedure for a Component Whose Failure Density is a Mixture of Exponential Failure Densities.* Naval Postgraduate School, Monterey. Unpublished thesis.

Vännman, K. (1976). Estimation based on order statistics from a Pareto distribution. *J. Amer. Statist. Assoc., 71,* 704-707.

Weibull, W. (1961). *Fatigue Testing and Analysis of Results.* Pergamon Press, New York.

Widder, D. V. (1946) *The LaPlace transform.* Princeton University Press, Princeton, New Jersey.

DA

FILM

7 –